

REQUEST FOR PARTICIPATION IN THE PATENT PROSECUTION HIGHWAY (PPH) PROGRAM BETWEEN THE JPO AND THE USPTO

Application No.:	10/692,222	First Named Inventor:	Hironobu Takagi
Filing Date:	10/23/2003	Attorney Docket No.:	JP920020152US1
Title of the Invention:	PROCESSING STRUCTURED/HIERARCHICAL CONTENT		

THIS REQUEST FOR PARTICIPATION IN THE PPH PROGRAM ALONG WITH THE REQUIRED DOCUMENTS MUST BE SUBMITTED VIA EFS-WEB. INFORMATION REGARDING EFS-WEB IS AVAILABLE AT [HTTP://WWW.USPTO.GOV/EBC/efs_help.html](http://www.uspto.gov/EBC/efs_help.html).

APPLICANT HEREBY REQUESTS PARTICIPATION IN THE PATENT PROSECUTION HIGHWAY (PPH) PROGRAM AND PETITIONS TO MAKE THE ABOVE-IDENTIFIED APPLICATION SPECIAL UNDER THE PPH PROGRAM.

The above-identified application (1) validly claims priority under 35 U.S.C. 119(a) and 37 CFR 1.55 to one or more corresponding JPO application(s) or to a PCT application that does not contain any priority claim, or (2) is a national stage entry of a PCT application that does not contain any priority claim.

The JPO/PCT application number(s) is/are: 2002-312331

The filing date of the JPO/PCT application(s) is/are: October 28, 2002

I. List of Required Documents:

a. A copy of the latest JPO office actions (other than "Decision to Grant a Patent") in the above-identified

JPO application(s)



Is attached.



No Office action from the JPO application is submitted since the JPO application was a first action allowance



Is available via Dossier Access System. Applicant hereby requests that the USPTO obtain these documents via the Dossier Access System.

*It is not necessary to submit a copy of the "Decision to Grant a Patent" and an English translation thereof.

b. A copy of all claims which were determined to be patentable by the JPO in the above-identified JPO application(s)



Is attached.



Is available via Dossier Access System. Applicant hereby requests that the USPTO obtain these documents via the Dossier Access System.

c. English translations of the documents in a. and b. above along with a statement that the English translations are accurate are attached (if the documents are not in the English language).

d. (1) An information disclosure statement listing the documents cited in the JPO office actions



Is attached.



Not applicable



Has already been filed in the above-identified U.S. application on _____

(2) Copies of all documents (except for U.S. patents or U.S. patent application publications)



Are attached.



Have already been filed in the above-identified U.S. application on _____

[Page 1 of 2]

This collection of information is required by 35 U.S.C. 119, 37 CFR 1.55, and 37 CFR 1.102(d). The information is required to obtain or retain a benefit by the public, which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 2 hours to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS.

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

**REQUEST FOR PARTICIPATION IN THE PATENT PROSECUTION HIGHWAY (PPH) PROGRAM BETWEEN
THE JPO AND THE USPTO**

(continued)

Application No.:	10/692,222	First Named Inventor:	Hironobu Takagi
------------------	------------	-----------------------	-----------------

II. Claims Correspondence Table:

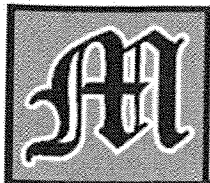
[illegible]

III. All the claims in the US application sufficiently correspond to the patentable/allowable claims in the JPO application.

IV. Payment of Fees:

The petition fee under 37 CFR 1.17(h) as required by 37 CFR 1.102(d) must be paid via EFS-Web (using credit card, authorization to charge a deposit account, or electronic funds transfer).

Signature	/Alexis Saenz/	Date	09/29/08
Name (Print/Typed)	Alexis Saenz	Registration Number	54570



TRANSLATOR CERTIFICATION

450 7th Ave | 6th Floor | New York, NY 10123 | Tel 212.643.8800 | Fax 212.643.0005 | www.mside.com

Morningside | Translations

The undersigned certifies that the attached translations are, to the best of my knowledge and belief, true and accurate translations from Japanese into English of:

- (1) the latest Notification of Reasons for Refusal by the Japanese Patent Office; and
- (2) all claims which were determined to be patentable by the Japanese Patent Office

from Japanese Patent Application No. 2002-312331, internal client docket number JP920020152JP1.

MORNINGSIDE EVALUATIONS AND CONSULTING


Shoko Sato

Date: September 9, 2008

拒絶理由通知書

特許出願の番号	特願 2 0 0 2 - 3 1 2 3 3 1
起案日	平成 1 8 年 8 月 1 5 日
特許庁審査官	深津 始 9 3 8 3 5 M 0 0
特許出願人復代理人	山崎 隆 様
適用条文	第 2 9 条柱書、第 3 6 条

この出願は、次の理由によって拒絶をすべきものである。これについて意見があれば、この通知書の発送の日から 3 か月以内に意見書を提出して下さい。

理 由

【理由 1】 この出願は、明細書及び図面の記載が下記の点で、特許法第 3 6 条第 4 項及び第 6 項第 2 号に規定する要件を満たしていない。

【理由 2】 この出願の下記の請求項に記載されたものは、下記の点で特許法第 2 9 条第 1 項柱書に規定する要件を満たしていないから、特許を受けることができない。

記

1. この出願には「構造化・階層化コンテンツ用処理装置」の主要構成に係る「構造化・階層化コンテンツが所定のマッチング・パターンとマッチするコンテンツ部分を含むか否かを判定し、」という判定を、当業者が実施をすることができる程度に記載されていない。

【0 0 0 7】などの記載より、判定が正であるときの所定の処理は、たとえば構造化・階層化コンテンツから一部を切り出す処理と認められる。しかし、切り出す部分を特定するために、マッチング・パターンを用いてどのような判定を行うのか、判定結果として何が得られるのか等、判定に関する事項が具体的に何ら開示されていない。

してみると、発明の詳細な説明は、当業者が実施をすることができる程度に明確かつ十分に記載したものではない。

2. 請求項 1 2 - 2 2 で各ステップにおける動作の主体が特定されておらず、たとえば「人間がマッチング・パターンを抽出しようとする構造化・階層化コンテンツ（以下、該構造化・階層化コンテンツを「ターゲット・コンテンツ」と言う。）におけるマッチング・パターンの抽出部分としてのターゲット・コンテンツ部分を含む範囲に係るターゲット・サブツリーを設定するターゲット・サブツリー設定ステップ」などと解釈できる。よって、請求項の記載は人間が各ステッ

プを実行する方法という「人間による方法」と解釈される。ここで、請求項 1 2－2 2に係る「人間による方法」とは、人間の手順すなわち人為的な取り決めであって、自然法則を利用したものではない。よって、請求項 1 2－2 2には自然法則を具体的に利用したものが記載されていないから、特許法第 2 条に定義される「発明」に該当しない。（第 2 9 条第 1 項柱書）

また、請求項 1 2－2 2で特許を受けようとする発明が、発明の詳細な説明にある「コンピュータ・ソフトウェアによる情報処理方法」という概念であるならば、請求項の記載は「人間による方法」とも解釈できる表現であり、「コンピュータ・ソフトウェアによる情報処理方法」とであると特定することができない表現であるから、請求項に係る記載は明確ではない。（第 3 6 条第 6 項第 2 号）

3. 【0 0 1 0】に係る関連出願の内容は次のものである。

【発明の属する技術分野】

本発明は、情報処理装置及びプログラムに関する。特に本発明は、表示装置上に表示するポインタの表示座標を、入力装置からの入力座標に基づいて定める情報処理装置及びプログラムに関する。

【従来の技術】

従来、デスクトップ型やノート型等のパーソナルコンピュータ、携帯情報端末、端末装置等の情報処理装置へ座標を入力する手段として、マウス、トラックボール、トラックポイント（商標）、タッチパネル、タッチパッド、又はタブレット等のポインティングデバイスによる入力装置が多用されている。この様な入力装置は、主に、表示装置上に表示された文字又は図形を指示するポインタの表示座標を指定又は変更するために使用される。

この様な入力装置は、主に人間の操作をデジタル値である入力座標に変換する。この変換の過程で、ノイズによる誤差、人間の指の微小な震え、操作を入力したアナログ値が 2 つのデジタル値の境界に位置することによるゆれ等の要因で、入力座標の下位ビットに変動が生じる。この結果、表示装置上のポインタが細かく振動する。

本願の技術内容と比べると、関連出願の番号に誤りがあると推定される。

拒絶の理由が新たに発見された場合には拒絶の理由が通知される。

先行技術文献調査結果の記録

- ・調査した分野 G 0 6 F 1 7 / 3 0
 D B 名
- J O I S
- ・先行技術文献 特開 2 0 0 2－2 4 5 0 6 8 号公報
 特開 2 0 0 0－2 2 7 9 9 6 号公報
 特開 2 0 0 1－2 5 9 8 4 6 号公報

福田健太郎、ほか、Webコンテンツ間の共通レイアウト自動解析，情報処理学会研究報告，2001年 5月25日，
2001巻、52号，7－14頁，2001-DD-28-2

この先行技術文献調査結果の記録は、拒絶理由を構成するものではない。

部長／代理	審査長／代理	審査官	審査官補
丹治 彰	深津 始		
8320	9383		

Notification of Reason for Refusal

Patent Application Number: Japanese Patent Application 2002-312331
Drafting Date: August 15, 2006
Examiner of JPO: Hajime Fukatsu 9383 5M00
Sub-agent of Applicant: Mr. Takashi Yamazaki
Applied Provision: Text of Patent Law Section 29, Patent Law Section 36

This application should be refused for the reason for refusal mentioned below. If the applicant has an opinion against this reason, such opinion needs to be submitted within three months from the dispatch date of this notification.

Reasons

[Reason 1] The descriptions of the specification and drawings in this subject application do not comply with the requirements under Patent Law Sections 36 (4) and (6) (ii) in the points mentioned below.

[Reason 2] The inventions relating to claims of the subject application mentioned below should not be granted a patent since they do not comply with the requirements under the text of Patent Law Section 29 (1) in the points mentioned below.

Note

1. This application is not described so as to enable a person with ordinary skills in the art pertaining to the present invention to implement a process “to make a determination about whether or not the structured/hierarchical content includes a content portion, which is matched with a predetermined matching pattern” relating to the main configuration of “structured/hierarchical content processing device”.

According to the description of [0007], it is perceived that the predetermined processing when the determination is positive is processing to clip a portion from the structured/hierarchical content. However, for the purpose of specifying the portion to be clipped, matters relating to a determination, such as a type of determination using a matching pattern or obtainment as a determination result, are not specifically disclosed at all.

Therefore, 'Detailed Description of the Invention' is not clearly and sufficiently described so as to enable a person with ordinary skills in the art pertaining to the present invention to implement the processing.

2. A subject of the action in each step is not specified in Claims 12 to 22, and for example, it can be interpreted, "a target subtree setting step of setting a target subtree relating to the scope including the target content portion as an extraction portion of the matching pattern in the structured/hierarchical content where a human tries to extract the matching pattern (hereafter, said structured/hierarchical content is referred to as "target content")". Therefore, the descriptions of claims are interpreted as "method by a human", which is a method to execute each step by a human. Herein, "the method by a human" relating to Claims 12 to 22 is a human procedure, i.e., an artificial arrangement, but not utilizes the natural laws. Therefore, because nothing specifically utilizing the natural laws is described in Claims 12 to 22, it does not correspond to "invention" defined in Patent Law Section 2 (text of Patent Law Section 29 (1)).

Further, if the inventions that are attempted to be granted a patent with Claims 12 to 22 are a concept, which is "information processing method by computer software", the description of the claims are expressions that can be interpreted as "method by a human", but they are not expressions that can be specified as "information processing method by computer software"; therefore, the descriptions relating to the claims are ambiguous (Patent Law Section 36 (6) (ii)).

3. The contents of the related application relating to [0010] are mentioned below.

[Technical Field of the Invention]

The present invention relates to an information processing apparatus and a program. Particularly, the present invention relates to an information processing apparatus and a program where display coordinates of a pointer displaying on the display unit are defined based upon input coordinates from the input device.

[Prior Art]

Conventionally, an input device by a mouse, a trackball, a track point (trademark), a touch panel, a touch pad or a pointing device, such as a tablet, is frequently used as input means of coordinates into an information processing apparatus, such as desktop or notebook personal computer, a personal digital assistance or a terminal unit. This type of input device is mainly used for designating or modifying display coordinates of the pointer indicating characters or figures displayed on the display unit.

In this type of input device, mainly operations by a human are converted into input coordinates, which are digital values. In this conversion process, a fluctuation occurs to a lower bit of the input coordinates due to a factor, such as an error by noises, minute trembling of human's fingers, vibration caused by a condition where an analog value entering the operation is located at a boundary of two digital values. As a result, the pointer on the display unit is minutely vibrated.

Compared to the technical contents of the subject application, it is assumed that the related application number has an error.

If a reason for refusal is newly discovered, the reason will be notified.

Record of Prior Art Document Search Results

- Searched technical field G06F 17/30
DB name
JOIS
- Prior art documents Japanese Unexamined Patent Application 2002-245068
Japanese Unexamined Patent Application 2000-227996
Japanese Unexamined Patent Application 2001-259846

Kentaro Fukuda, et al., Common layout automatic analysis
between Web contents, Research Report of Information
Processing Society of Japan, May 25, 2001, Vol. 2001, No.
52, pp. 7 – 14, 2001-DD-28-2

The record of prior art document search results is not a component of reasons for
refusal.

<u>Manager/Deputy</u>	<u>Chief Examiner/Deputy</u>	<u>Examiner</u>	<u>Assistant Examiner</u>
	Akira Tanji	Hajime Fukatsu	
	8 8 4 1	9 1 9 3	

【特許請求の範囲】

【請求項1】

ネットワークを介して配信される構造化・階層化コンテンツの記述と、マッチング・パターンの記述とに基づき、マッチング・パターンの記述に対応する記述が前記構造化・階層化コンテンツの記述に含まれているか否かを判定し、該判定が正であれば、マッチング・パターンの記述に対応する構造化・階層化コンテンツ側記述部分をマッチング・コンテンツ部分とし、該マッチング・コンテンツ部分に対して所定の処理を行う構造化・階層化コンテンツ用処理装置であって、

マッチング・パターンを抽出しようとする構造化・階層化コンテンツ（以下、該構造化・階層化コンテンツを「ターゲット・コンテンツ」と言う。）におけるマッチング・パターンの抽出部分としてのターゲット・コンテンツ部分を含む範囲に係るターゲット・サブツリーを設定するターゲット・サブツリー設定手段、

前記ターゲット・コンテンツに対する過去の複数の構造化・階層化コンテンツを選択し前記ターゲット・コンテンツに係るターゲット・サブツリーと過去の各構造化・階層化コンテンツに係るツリーとを対照してターゲット・サブツリーの各ノードの出現態様を検出する出現態様検出手段、

過去の複数の構造化・階層化コンテンツに基づいて該ターゲット・サブツリーにおける各ノードについての出現態様の出現頻度に係る統計情報を生成する統計情報生成手段、

前記出現態様検出結果及び前記統計情報に基づいてターゲット・サブツリーの各ノードを分類する分類手段、及び

該分類に基づいて前記ターゲット・コンテンツ部分についてのマッチング・パターンを生成するマッチング・パターン生成手段、

を有していることを特徴とする構造化・階層化コンテンツ用処理装置。

【請求項2】

前記所定の処理とは、該構造化・階層化コンテンツのコンテンツ部分への関連情報の関連付けであることを特徴とする請求項1記載の構造化・階層化コンテンツ用処理装置。

【請求項3】

前記関連情報はアノテーションを含むことを特徴とする請求項2記載の構造化・階層化コンテンツ用処理装置。

【請求項4】

前記所定の処理とは、構造化・階層化コンテンツのコンテンツ部分を他の構造化・階層化コンテンツに利用するために該構造化・階層化コンテンツの該コンテンツ部分をコピーする処理であることを特徴とする請求項1記載の構造化・階層化コンテンツ用処理装置。

【請求項5】

構造化・階層化コンテンツとはウェブ・コンテンツであることを特徴とする請求項1記載の構造化・階層化コンテンツ用処理装置。

【請求項6】

ターゲット・サブツリーのノードを、定常ノード、更新ノード及び付加ノードに分類する前記分類手段を有していることを特徴とする請求項1記載の構造化・階層化コンテンツ用処理装置。

【請求項7】

検出する前記出現態様として、(N1) 被検出ノードがターゲット・コンテンツ部分及び対照構造化・階層化コンテンツの両方に出現しその内容が相互に同一となっている出現態様、及び(N2) 被検出ノードがターゲット・コンテンツ部分及び対照構造化・階層化コンテンツの両方に出現しその内容が相互に異なっている出現態様を含む前記出現態様検出手段、及び

統計情報により(N1)の出現態様による出現頻度が第1の閾値以上であると判明したノードは定常ノードに分類し、統計情報により(N2)の出現態様による出現頻度が第2の閾値以上であると判明したノードは更新ノードに分類し、定常ノード及び更新ノード以外のノードは付加ノードに分類する前記分類手段、

を有していることを特徴とする請求項6記載の構造化・階層化コンテンツ用処理装置。

【請求項8】

前記マッチング・パターン生成手段は、

定常ノード、更新ノード及び付加ノードの分類に基づいてターゲット・サブツリーにおける繰り返し部分を検出する繰り返し部分検出手段、及び

該繰り返し部分の存在情報を含む前記マッチング・パターンを生成する繰り返し情報付きマッチング・パターン生成手段、

を有していることを特徴とする請求項6記載の構造化・階層化コンテンツ用処理装置。

【請求項9】

前記分類手段は、

イメージに係るノードについて、該ノードが空白領域を確保するためのスペーサ用イメージに係るノードであるか否かを検出するスペーサ用イメージ検出手段、

イメージに係るノードについて、該ノードが繰り返して同一サイズで複数個使用されるビュレット・イメージに係るノードであるか否かを検出するビュレット・イメージ検出手段、
スペーサ用イメージに係るノードは付加ノードと分類する第1の分類付け手段、

ビュレット・イメージに係るノード同士は、その表示内容が異なっても定常ノード、更新ノード又は付加ノードの同一分類に割り当てる第2の分類付け手段、

を有していることを特徴とする請求項8記載の構造化・階層化コンテンツ用処理装置。

【請求項10】

ターゲット・コンテンツに対する過去の構造化・階層化コンテンツが存在しない場合には、過去の各構造化・階層化コンテンツの代わりに該ターゲット・コンテンツに対する複数個の隣接構造化・階層化コンテンツを選択しターゲット・コンテンツに係るターゲット・サ

ブツリーと各隣接構造化・階層化コンテンツに係るツリーと対照する前記対照手段、
を有していることを特徴とする請求項 1 記載の構造化・階層化コンテンツ用処理装置。

【請求項 1 1】

ネットワークを介して配信される構造化・階層化コンテンツの記述と、マッチング・パターンの記述とに基づき、マッチング・パターンの記述に対応する記述が前記構造化・階層化コンテンツの記述に含まれているか否かを判定し、該判定が正であれば、マッチング・パターンの記述に対応する構造化・階層化コンテンツ側記述部分をマッチング・コンテンツ部分とし、該マッチング・コンテンツ部分に対して所定の処理を行う構造化・階層化コンテンツ用処理装置であって、

マッチング・パターンを抽出しようとする構造化・階層化コンテンツ（以下、該構造化・階層化コンテンツを「ターゲット・コンテンツ」と言う。）におけるマッチング・パターンの抽出部分としてのターゲット・コンテンツ部分を含む範囲に係るターゲット・サブツリーを設定するターゲット・サブツリー設定手段、

前記ターゲット・コンテンツに対する複数の隣接構造化・階層化コンテンツを選択し前記ターゲット・コンテンツに係るターゲット・サブツリーと各隣接構造化・階層化コンテンツに係るツリーとを対照してターゲット・サブツリーの各ノードの出現態様を検出する出現態様検出手段、

過去の複数の構造化・階層化コンテンツに基づいて該ターゲット・サブツリーにおける各ノードについての出現態様の出現頻度に係る統計情報を生成する統計情報生成手段、

前記出現態様検出結果及び前記統計情報に基づいてターゲット・サブツリーの各ノードを分類する分類手段、及び

該分類に基づいて前記ターゲット・コンテンツ部分についてのマッチング・パターンを生成するマッチング・パターン生成手段、

を有していることを特徴とする構造化・階層化コンテンツ用処理装置。

【請求項 1 2】

ネットワークを介して配信される構造化・階層化コンテンツが所定のマッチング・パターンとマッチするコンテンツ部分を含むか否かを判定し、該判定が正であれば該構造化・階層化コンテンツについて所定の処理を行う構造化・階層化コンテンツ用処理方法であって、コンピュータが、マッチング・パターンを抽出しようとする構造化・階層化コンテンツ（以下、該構造化・階層化コンテンツを「ターゲット・コンテンツ」と言う。）におけるマッチング・パターンの抽出部分としてのターゲット・コンテンツ部分を含む範囲に係るターゲット・サブツリーを設定するターゲット・サブツリー設定ステップ、

前記コンピュータが、前記ターゲット・コンテンツに対する過去の複数の構造化・階層化コンテンツを選択し前記ターゲット・コンテンツに係るターゲット・サブツリーと過去の各構造化・階層化コンテンツに係るツリーとを対照してターゲット・サブツリーの各ノードの出現態様を検出する出現態様検出ステップ、

前記コンピュータが、過去の複数の構造化・階層化コンテンツに基づいて該ターゲット・サブツリーにおける各ノードについての出現態様の出現頻度に係る統計情報を生成する統計情報生成ステップ、

前記コンピュータが、前記出現態様検出結果及び前記統計情報に基づいてターゲット・サブツリーの各ノードを分類する分類ステップ、及び

前記コンピュータが、該分類に基づいて前記ターゲット・コンテンツ部分についてのマッチング・パターンを生成するマッチング・パターン生成ステップ、

を有していることを特徴とする構造化・階層化コンテンツ用処理方法。

【請求項 13】

前記所定の処理とは、該構造化・階層化コンテンツのコンテンツ部分への関連情報の関連付けであることを特徴とする請求項 12 記載の構造化・階層化コンテンツ用処理方法。

【請求項 14】

前記関連情報はアノテーションを含むことを特徴とする請求項 13 記載の構造化・階層化コンテンツ用処理方法。

【請求項 15】

前記所定の処理とは、構造化・階層化コンテンツのコンテンツ部分を他の構造化・階層化コンテンツに利用するために該構造化・階層化コンテンツの該コンテンツ部分をコピーする処理であることを特徴とする請求項 12 記載の構造化・階層化コンテンツ用処理方法。

【請求項 16】

構造化・階層化コンテンツとはウェブ・コンテンツであることを特徴とする請求項 12 記載の構造化・階層化コンテンツ用処理方法。

【請求項 17】

前記コンピュータが、ターゲット・サブツリーのノードを、定常ノード、更新ノード及び付加ノードに分類する前記分類ステップを有していることを特徴とする請求項 12 記載の構造化・階層化コンテンツ用処理方法。

【請求項 18】

検出する前記出現態様として、(N1) 被検出ノードがターゲット・コンテンツ部分及び対照構造化・階層化コンテンツの両方に出現しその内容が相互に同一となっている出現態様、及び (N2) 被検出ノードがターゲット・コンテンツ部分及び対照構造化・階層化コンテンツの両方に出現しその内容が相互に異なっている出現態様を含む前記出現態様検出ステップ、及び

前記コンピュータが、統計情報により (N1) の出現態様による出現頻度が第 1 の閾値以上であると判明したノードは定常ノードに分類し、統計情報により (N2) の出現態様による出現頻度が第 2 の閾値以上であると判明したノードは更新ノードに分類し、定常ノード及び更新ノード以外のノードは付加ノードに分類する前記分類ステップ、

を有していることを特徴とする請求項 17 記載の構造化・階層化コンテンツ用処理方法。

【請求項 19】

前記マッチング・パターン生成ステップは、
前記コンピュータが、定常ノード、更新ノード及び付加ノードの分類に基づいてターゲット・サブツリーにおける繰り返し部分を検出する繰り返し部分検出ステップ、及び
前記コンピュータが、該繰り返し部分の存在情報を含む前記マッチング・パターンを生成する繰り返し情報付きマッチング・パターン生成ステップ、
を有していることを特徴とする請求項 17 記載の構造化・階層化コンテンツ用処理方法。

【請求項 20】

前記分類ステップは、
前記コンピュータが、イメージに係るノードについて、該ノードが空白領域を確保するためのスペーサ用イメージに係るノードであるか否かを検出するスペーサ用イメージ検出ステップ、
前記コンピュータが、イメージに係るノードについて、該ノードが繰り返して同一サイズで複数個使用されるビュレット・イメージに係るノードであるか否かを検出するビュレット・イメージ検出ステップ、
前記コンピュータが、スペーサ用イメージに係るノードは付加ノードと分類する第 1 の分類付けステップ、
前記コンピュータが、ビュレット・イメージに係るノード同士は、その表示内容が異なっても定常ノード、更新ノード又は付加ノードの同一分類に割り当てる第 2 の分類付けステップ、
を有していることを特徴とする請求項 19 記載の構造化・階層化コンテンツ用処理方法。

【請求項 21】

前記コンピュータが、ターゲット・コンテンツに対する過去の構造化・階層化コンテンツが存在しない場合には、過去の各構造化・階層化コンテンツの代わりに該ターゲット・コンテンツに対する複数の隣接構造化・階層化コンテンツを選択しターゲット・コンテンツに係るターゲット・サブツリーと各隣接構造化・階層化コンテンツに係るツリーと対照する前記対照ステップ、
を有していることを特徴とする請求項 12 記載の構造化・階層化コンテンツ用処理方法。

【請求項 22】

ネットワークを介して配信される構造化・階層化コンテンツの記述と、マッチング・パターンの記述とに基づき、マッチング・パターンの記述に対応する記述が前記構造化・階層化コンテンツの記述に含まれているか否かを判定し、該判定が正であれば、マッチング・パターンの記述に対応する構造化・階層化コンテンツ側記述部分をマッチング・コンテンツ部分とし、該マッチング・コンテンツ部分に対して所定の処理を行う構造化・階層化コンテンツ用処理方法であって、
コンピュータが、マッチング・パターンを抽出しようとする構造化・階層化コンテンツ（以

下、該構造化・階層化コンテンツを「ターゲット・コンテンツ」と言う。)におけるマッチング・パターンの抽出部分としてのターゲット・コンテンツ部分を含む範囲に係るターゲット・サブツリーを設定するターゲット・サブツリー設定ステップ、
前記コンピュータが、前記ターゲット・コンテンツに対する複数の隣接構造化・階層化コンテンツを選択し前記ターゲット・コンテンツに係るターゲット・サブツリーと各隣接構造化・階層化コンテンツに係るツリーとを対照してターゲット・サブツリーの各ノードの出現態様を検出する出現態様検出ステップ、
前記コンピュータが、過去の複数の構造化・階層化コンテンツに基づいて該ターゲット・サブツリーにおける各ノードについての出現態様の出現頻度に係る統計情報を生成する統計情報生成ステップ、
前記コンピュータが、前記出現態様検出結果及び前記統計情報に基づいてターゲット・サブツリーの各ノードを分類する分類ステップ、及び
前記コンピュータが、該分類に基づいて前記ターゲット・コンテンツ部分についてのマッチング・パターンを生成するマッチング・パターン生成ステップ、
を有していることを特徴とする構造化・階層化コンテンツ用処理方法。

【請求項 23】

請求項 12～22 のいずれかに記載の構造化・階層化コンテンツ用処理方法の各ステップをコンピュータに実行させるプログラム。

CLAIMS

What is claimed is:

1. A processing apparatus for a structured/hierarchical content, based upon a description of structured/hierarchical content to be distributed via a network and a description of matching pattern, making a determination about whether or not a description corresponding to the description of matching pattern is included in the description of structured/hierarchical content, and regarding the description portion of the structured/hierarchical content corresponding to the description of the matching pattern as a matching content portion if a result of the determination is positive, and performing predetermined processing for the structured/hierarchical content, comprising:

target subtree setting means for setting a target subtree relating to a range including a target content portion as an extracted portion of the matching pattern in the structure/hierarchical content (hereafter, referred to as a “target content”) from which the matching pattern is to be extracted;

occurrence mode defecting means for detecting an occurrence mode of each node of the target subtree by selecting a plurality of past structured/hierarchical contents with respect to the target content and collating the target subtree relating to the target content with a tree relating to each of the past structure/hierarchical contents;

statistical information generating means for generating statistical information concerning an occurrence frequency of the occurrence mode of each node in the target subtree based on the plurality of past structured/hierarchical contents;

classifying means for performing classification of each node of the target subtree

based on a result of detecting the occurrence mode and the statistical information; and
matching pattern generating means for generating the matching pattern for the
target content portion based on the classification.

2. The processing apparatus for a structured/hierarchical content according to claim
1, wherein

the predetermined processing is to associated related information with the content
portion of the structured/hierarchical content.

3. The processing apparatus for a structured/hierarchical content according to claim
2, wherein

the related information includes an annotation.

4. The processing apparatus for a structured/hierarchical content according to claim
1, wherein

the predetermined processing is processing for copying the content portion of the
structured/hierarchical content for a purpose of utilizing the content portion of the
structured/hierarchical content for another structured/hierarchical content.

5. The processing apparatus for a structured/hierarchical content according to claim
1, wherein

the structured/hierarchical content is a Web content.

6. The processing apparatus for a structured/hierarchical content according to claim

1, comprising the classifying means for classifying nodes of the target subtree into stationary nodes, updated nodes and additional nodes.

7. The processing apparatus for a structured/hierarchical content according to claim 6, comprising:

the occurrence node detecting means for classifying, as the occurrence mode to be detected, (N1) an occurrence mode where detected nodes occur in both of the target content portion and structured/hierarchical contents collated therewith and contents thereof are mutually identical, and (N2) an occurrence mode where the detected nodes occur in both of the target content portion and the structured/hierarchical contents collated therewith and the contents thereof are mutually different, and

the classifying means for classifying, into the stationary nodes, nodes of which occurrence frequency of the occurrence mode (N1) is determined to be equal to/more than a first threshold value by the statistical information, and for classifying, into the updated nodes, nodes, of which occurrence frequency of the occurrence node (N2) is determined to be equal to/more than a second threshold value by the statistical information, and classifying, into the additional nodes, nodes other than the stationary nodes and the updated nodes.

8. The processing apparatus for a structured/hierarchical content according to claim 6, wherein

the matching pattern generating means includes:

repeated portion detecting means for detecting a repeated portion in the target subtree based on the classification into the stationary nodes, the updated nodes and the

2002-312331

additional nodes; and

repeated information-added matching pattern generating means for generating the matching pattern including presence information of the repeated portion.

9. The processing apparatus for a structured/hierarchical content according to claim 8, wherein

the classifying means includes:

spacer image detecting means for detecting whether or not a node relating to an image is a node relating to a spacer image for ensuring a blank region;

bullet image detecting means for detecting whether or not the node relating to the image is a node relating to a plurality of bullet images used repeatedly in a same size;

first classifying means for classifying the node relating to the spacer image into the additional nodes;

second classifying means for allocating nodes relating to the bullet image into a same classification among classifications of the stationary nodes, updated nodes and additional nodes even if display contents of the nodes are mutually different.

10. The processing apparatus for a structured/hierarchical content according to claim 1, comprising:

collating means for selecting, instead of each of past structured/hierarchical contents, a plurality of structured/hierarchical contents with regard to said target contents in the case that no past structured/hierarchical content with regard to the target content exists, and for collating the target subtree relating to the target content with the trees relating to each structured/hierarchical content.

11. A processing apparatus for a structured/hierarchical content, based upon a description of structured/hierarchical content to be distributed via a network and a description of matching pattern, making a determination about whether or not a description corresponding to the description of matching pattern is included in the description of structured/hierarchical content, and regarding the description portion of the structured/hierarchical content corresponding to the description of the matching pattern as a matching content portion if a result of the determination is positive, and performing predetermined processing for the structured/hierarchical content, comprising:

target subtree setting means for setting a target subtree relating to a range including a target content portion as an extracted portion of the matching pattern in the structure/hierarchical content (hereafter, referred to as a “target content”) from which the matching pattern is to be extracted;

occurrence mode defecting means for detecting an occurrence mode of each node of the target subtree by selecting a plurality of adjacent structured/hierarchical contents with respect to the target content and collating the target subtree relating to the target content with a tree relating to each of the adjacent structure/hierarchical contents;

statistical information generating means for generating statistical information concerning an occurrence frequency of the occurrence mode of each node in the target subtree based on the plurality of past structured/hierarchical contents;

classifying means for performing classification of each node of the target subtree based on a result of detecting the occurrence mode and the statistical information; and

matching pattern generating means for generating the matching pattern for the target content portion based on the classification.

12. A processing method for a structured/hierarchical content for making a determination about whether or not structured/hierarchical content delivered through a network includes a content portion matched with a predetermined matching pattern, and for performing predetermined processing for the structured/hierarchical content if a result of the determination is positive, comprising:

a target subtree setting step of setting a target subtree relating to a range including a target content portion as an extracted portion of the matching pattern in the structure/hierarchical content (hereafter, referred to as a “target content”) from which the matching pattern is to be extracted by the computer;

an occurrence mode defecting step of detecting an occurrence mode of each node of the target subtree by selecting a plurality of past structured/hierarchical contents with respect to the target content and collating the target subtree relating to the target content with a tree relating to each of the past structure/hierarchical contents by the computer;

a statistical information generating step of generating statistical information concerning an occurrence frequency of the occurrence mode of each node in the target subtree based on the plurality of past structured/hierarchical contents by the computer;

a classifying step of performing classification of each node of the target subtree based on a result of detecting the occurrence mode and the statistical information; and

matching pattern generating step of generating the matching pattern for the target content portion based on the classification by the computer.

13. The processing method for a structured/hierarchical content according to claim 12, wherein

the predetermined processing is to associate related information with the content portion of the structured/hierarchical content.

14. The processing method for a structured/hierarchical content according to claim 13, wherein

the related information includes an annotation.

15. The processing method for a structured/hierarchical content according to claim 12, wherein

the predetermined processing is processing for copying the content portion of the structured/hierarchical content for a purpose of utilizing the content portion of the structured/hierarchical content for another structured/hierarchical content.

16. The processing method for a structured/hierarchical content according to claim 12, wherein

the structured/hierarchical content is a Web content.

17. The processing method for a structured/hierarchical content according to claim 12, comprising the classifying step where the computer classifies nodes of the target subtree into stationary nodes, updated nodes and additional nodes.

18. The processing method for a structured/hierarchical content according to claim 17, comprising:

the occurrence node detecting step of classifying, as the occurrence mode to be

detected, (N1) an occurrence mode where detected nodes occur in both of the target content portion and structured/hierarchical contents collated therewith and contents thereof are mutually identical, and (N2) an occurrence mode where the detected nodes occur in both of the target content portion and the structured/hierarchical contents collated therewith and the contents thereof are mutually different, and

the classifying step of classifying, into the stationary nodes, nodes of which occurrence frequency of the occurrence mode (N1) is determined to be equal to/more than a first threshold value by the statistical information, and for classifying, into the updated nodes, nodes, of which occurrence frequency of the occurrence node (N2) is determined to be equal to/more than a second threshold value by the statistical information, and classifying, into the additional nodes, nodes other than the stationary nodes and the updated nodes, by the computer.

19. The processing method for a structured/hierarchical content according to claim 17, wherein

the matching pattern generating step includes:

a repeated portion detecting step of detecting a repeated portion in the target subtree based on the classification into the stationary nodes, the updated nodes and the additional nodes by the computer; and

a repeated information-added matching pattern generating step of generating the matching pattern including presence information of the repeated portion by the computer.

20. The processing method for a structured/hierarchical content according to claim 19, wherein

2002-312331

the classifying step includes:

a spacer image detecting step of detecting whether or not a node relating to an image is a node relating to a spacer image for ensuring a blank region by the computer;

a bullet image detecting step of detecting whether or not the node relating to the image is a node relating to a plurality of bullet images used repeatedly in a same size by the computer;

a first classifying step of classifying the node relating to the spacer image into the additional nodes by the computer;

a second classifying step of allocating nodes relating to the bullet image into a same classification among classifications of the stationary nodes, updated nodes and additional nodes even if display contents of the nodes are mutually different by the computer.

21. The processing apparatus for a structured/hierarchical content according to claim 12, comprising:

a collating step of selecting, instead of each of past structured/hierarchical contents, a plurality of structured/hierarchical contents with regard to said target contents in the case that no past structured/hierarchical content with regard to the target content exists, and for collating the target subtree relating to the target content with the trees relating to each structured/hierarchical content by the computer.

22. A processing method for a structured/hierarchical content, based upon a description of structured/hierarchical content to be distributed via a network and a description of matching pattern, making a determination about whether or not a

description corresponding to the description of matching pattern is included in the description of structured/hierarchical content, and regarding the description portion of the structured/hierarchical content corresponding to the description of the matching pattern as a matching content portion if a result of the determination is positive, and performing predetermined processing for the structured/hierarchical content, comprising: a target subtree setting step of setting a target subtree relating to a range including a target content portion as an extracted portion of the matching pattern in the structure/hierarchical content (hereafter, referred to as a “target content”) from which the matching pattern is to be extracted;

an occurrence mode detecting step of detecting an occurrence mode of each node of the target subtree by selecting a plurality of adjacent structured/hierarchical contents with respect to the target content and collating the target subtree relating to the target content with a tree relating to each of the adjacent structure/hierarchical contents;

a statistical information generating step of generating statistical information concerning an occurrence frequency of the occurrence mode of each node in the target subtree based on the plurality of past structured/hierarchical contents;

a classifying step of performing classification of each node of the target subtree based on a result of detecting the occurrence mode and the statistical information; and

a matching pattern generating step of generating the matching pattern for the target content portion based on the classification.

23. A program causing a computer to execute each step of the processing method for a structured/hierarchical content according to any of claims 12 to 22.